Detecting Tweet-Based Sentiment Polarity of Plastic Surgery Treatment

MARVI S. JOKHIO*, NAEEM AHMED MAHOTO**, SALAHUDDIN JOKHIO***, AND MUHAMMAD SHABAN JOKHIO****

RECEIVED ON 01.01.2015 ACCEPTED ON 28.05.2015

ABSTRACT

Sentiment analysis is a growing research these days. Many companies perform this analysis on public opinions to get a general idea about any product or service. This paper presents a novel approach to get views or comments of Twitter users about plastic surgery treatments. The proposed approach uses machine-learning technique embedded with the naïve Bayesian classifier to assign polarities (i.e. positive, negative or neutral) to the tweets, collected from "Twitter micro-blogging website". The accuracy of the obtained results has been validated using precision, recall and F-score measures. It has been observed from 25000 tweets dataset that people tend to have positive as well as substantial negative opinions regarding particular treatments. The experimental results show the effectiveness of the proposed approach.

Key Words: Sentiment Analysis, Tweets, Plastic Surgery, Classification.

1. INTRODUCTION

The individuals who want to get the shape of particular part of their body changed, due to medical or any cosmetic reasons go through the plastic surgery treatment. According to the statistics by American Society of Plastic Surgeons for 2013 [1], 3% rise has been observed in plastic surgery as compare to 2012. In year 2013, 15.1 million plastic surgery procedures were performed only in the United States. After treatment, some people are satisfied and others are disappointed. Most of the people nowadays go through the related reviews over the Internet before getting medical treatment. This is particularly true for the plastic surgery. Although, several documents can be accessed over the WWW (World Wide Web). However, the reliability and quality of information of such documents is vague. The information available at websites, blogs and documents may not be accurate or can be misleading. Finding an appropriate and relevant information to one's query from immense collection of data is a challenging is difficult task. In particular, it is a relatively harder to get sentiments or opinions of the online users about certain topic or issue. Therefore, a potential method is needed to determine public's general tendency in the form of Yes or No answers (i.e. positive or negative sentiment) towards certain query related to a particular topic. Sentiment analysis can be an appropriate solution in this scenario.

* Lecturer, Government Girls College, Qasimabad, Hyderabad.
** Assistant Professor, Department of Software Engineering, Mehran University of Engineering & Technology, Jamshoro.
*** Lecturer, Department of Software Engineering, Mehran University of Engineering & Technology, Jamshoro.
**** Assistant Professor, Department of Computer Systems Engineering, Mehran University of Engineering & Technology, Jamshoro.

Sentiment analysis is widely used in several applications. For instance, customer feedbacks on products or services, public opinion on government's decisions, public sentiments on particular event and news alerts [2]. In past, Twitter has been used to perform the users' sentiment analysis during elections [3-4] and used to predict stock market [5-7].

Besides, sentiment analysis has also effect in improving medical facilities and healthcare [8]. Analysis of healthcare opinions is an emerging trend that can facilitate health care organizations to take competitive advantages in understanding and improving the patient's practices. This can lead towards betterment of medical services. The patients provide their feedback and sentiments about the treatments on social media. Whereas, the other patients who may also plan for similar treatments can beneficially use this feedback and patient's sentiments for future considerations. Further, healthcare agencies may also get first-hand feedback to improve their services based on patient's sentiments.

This paper presents a novel approach to get the summarized results of user opinions about the plastic surgery treatment. The main focus of the paper is to discover overall sentiment polarity about the considered treatment using Twitter data as the corpus. In particular, a classifier is designed to assign a class of sentiment polarities (i.e. positive, negative or neutral) to individual tweets. The experiment has been conducted on real dataset collected from Twitter micro-blogging website. The results have been evaluated using performance matrices such as accuracy, precision, recall and F-measure. The extracted information shows the users' sentiments towards different plastic surgery treatments. The obtained results are quite appealing, which represent the effectiveness of the proposed approach.

The remaining paper is organized as follows. Section 2 reports related research; Section 3 thoroughly covers the proposed approach. Section 4 presents experimental results. Finally, Section 5 and 6 gives the conclusion and proposed future work respectively.

2. RELATED RESEARCH

There has been a growing interest in the field of sentiment analysis for past many years [9]. Several research experiments have been conducted in this domain. For instance, sentiment classification of tweets (i.e. positive, negative and neutral classes) is reported in [8] to measure the degree of concern regarding the public health as general topic. They focused on the classification of sentiments in order to measure the DoC (Degree of Concern) of the twitter users. A two-step novel algorithm is also proposed to automatically classify the tweets in personal and negative sentiments [10]. Discusses the facts and figures about Autism Disorder using twitter as a data source through text mining. The study in [10] uses several dictionaries, and different approaches for sentiment analysis.

Ofek, et. al. [11], proposed an automated approach for performing the sentiment analysis in an online cancer survivor community. The sentiment analysis was performed using two approaches. Both approaches use machine learning and were tested on one dataset. But, this work has utilized features derived from a dynamic sentiment lexicon, whereas the previous work was using a general sentiment lexicon to extract the features.

Sentiment analysis about the latest Hollywood movies is presented in [12]. Sentiment analysis was performed over the population of the different parts of world. Sentiments were classified as positive, negative and cognitive statements. Prediction on results of general elections 2013 in Pakistan was made by using data mining from the Twitter data [13]. An approach to take into consideration the role of verbs, being the most important aspect in social issues analysis, is proposed in [14]. Study in [15] addresses the sentimental analysis about Thailand's tourism during early part of 2010. This paper presents detecting sentiments of Twitter users about plastic surgery treatments.

3. SENTIMENT DETECTION-PROPOSED APPROACH

To get users opinions, Twitter can be considered as a rich repository. Tweets are considered as good source for information extraction because of its limited size of text, global sheer magnitude of data and easiness to collect in vast amount. The proposed approach for detecting sentiments for plastic surgery treatment exploits tweets as the dataset. Fig. 1 presents the block diagram of the approach.

The detailed methodology to get information from tweets is described in the subsequent sections.

3.1 Information Extraction

This section reports the mechanism of deriving knowledge from social media.

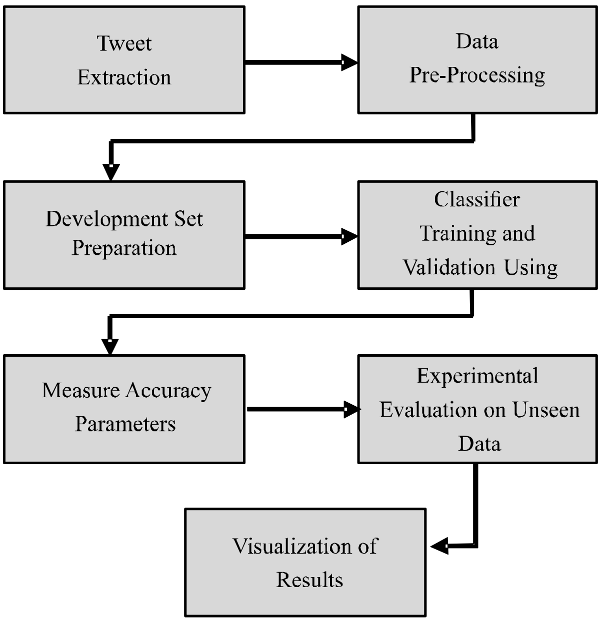

FIG. 1. BLOCK DIAGRAM OF THE PROPOSED APPROACH

***Extracting Tweets*:** Tweets are extracted from popular micro-blogging website Twitter using its public stream API (Application Programming Interface) [16], and R utility. The packages which are utilized in R language for this purpose of sentiment analysis are library(twitteR), library(plyr). Total 48 different keywords related to different plastic surgery have been used to extract the tweets. To select the appropriate keywords for the treatments, American Society of Plastic Surgeons website has been considered as reference. For every search query, maximum 1500 tweets are extracted and minimum 25 in some keywords also have been observed.

For example:

tweets<-searchTwitter("liposuction", cainfo="cacert.pem", n=1500)

Query searches maximum possible available tweets regarding liposuction procedure. Here, n=1500 represents the total number of 1500 tweets to be extracted.

Prepare the Text for Sentiment Analysis: The collected tweets are then preprocessed in order to get the desired textual information for plastic surgery treatment. The casual nature of twitter made this necessary to clean and filter the text like:

- Removed hash tags
- Removed RT retweets
- Removed @usernames
- Removed http:// urls
- Removed unnecessary spaces, symbols and punctuations.

To achieve this, gsub() method of R language has been used, which matches text in terms of regular expressions and manipulates as directed.

3.2 Sentiment Analysis using Machine Learning Approach

Different approaches can be used to identify text as positive or negative sentiment, such as lexicon-based and machine-learning approaches. In the proposed approach, a supervised machine learning technique has been used to classify the tweets of plastic surgery. The steps involved in the proposed approach during the supervised machine learning technique are described as follows:

- Building Training dataset
- Training the classifier
- Accuracy

Building Training Dataset: The Accuracy of supervised machine learning technique depends on size of training dataset. A training set of 3000 tweets regarding plastic surgery has been built. This training set is labeled through semi-automated method and each tweet is assigned a SO (Sentiment Orientation), which is numerical score defining strength of someone's personal belief, attitude or feelings.

To count the positive and negative words in a tweet Hu and Liu's English opinion lexicon [17] is used, which contains 2006 collection of positive and 4785 negative words. The text in a tweet is tokenized and each token is matched with positive and negative lexicon. The words that are matched in positive dictionary of lexicon are separated from that of negative matched ones. Finally, the orientation is calculated usingJaffrey Breen's approach [18]. Jaffrey Breen used this approach for scoring tweets regarding airline top consumer satisfaction [18].

Depending on value of SO ranging from -5 to +5 the corresponding tweet is labeled as:

- Positive, if SO > 0
- Negative, if SO< 0
- Neutral, if SO = 0

Table 1 shows some sample tweet records of the training data set.

Training the Classifier: In order to carry out the task of training, the Naïve Baye's classifier, a probabilistic model based on Baye's theorem, has been used as a means to classify the tweets. The reason behind using Naïve Baye's classifier is its potential efficiency in text classification, since it possesses linear and optimal training and testing time complexity during the data scanning [19].

TABLE 1.TRAINSET SAMPLE TWEETS

Tweet	SO Value	Polarit
Ugh That the worst But I cant even imagine how much more it would hurt after facial surgery	-3	Negative
If tattoos are a sin plastic surgery for aesthetic reasons must be a sin as well	-1	
Botulinum toxin is so dangerous that 2 kilograms would kill every human on earth	-2	
I am keen to get body contouring	1	Positive
Whats the difference between silicone and saline implants	0	Neutral
Key things to know before having a facial cosmetic surgery performed	0	
Her facial reconstruction surgery went great	3	Positive
Congratulations to Kerisha Mark and her amazing transformation and incredible courage Doesnt she look fabulous	5	
In my chin surgery I prefer my chin like that instead of my huge one I have the now	3	

$$P(c/t)=[P(c)P(t/c)]/P(t) \quad (1)$$

Here c represents a class and t represents text to be classified.

$$c = \arg\max_c P(c)\prod_{i=1}^{n} P(x_i|c) \quad (2)$$

Where c is the class, positive or negative, and xi is a word in a sentence.

The cross validation technique is used to train the classifier. The whole training set was divided into four parts and every time in training progression classifier is trained with diverse set of train data and test data to improve the overall accuracy. The WEKA tool has been used to train the naïve bayes classifier.

Accuracy: The accuracy of the classifier after 4 training folds resulted improved outcomes as shown in contingency Table 2.

Table 2 shows a total of 1886 tweets have been used under final train fold where number of predicted positives are 1666, which is very much close to the actual positives 1661. Similarly, values of predicted negatives 220 and actual negatives 225 are also closer to each other. Hence, the overall accuracy is much better, after 4 training folds.

Table 3 reports different performance measures. The 85.73% accuracy has been gained with misclassification rate much lower up to 14.2%. This shows that the final results of the classifier would be quite reasonable for considerations. The subsequent section reports the experiments carried out using this trained classifier.

It can be seen in Table 3, that the False Positive rate is very high, i.e. 60.8%. This is due to the dataset containing the tweets regarding advertisements, promotions and offers. Such tweets are written using words which are positive, but actually those tweets don't reflect sentiments of an individual.

3.3 Experiment on Unseen Dataset

The trained classifier has been run on an unseen dataset containing approximately 25000 extracted tweets of plastic surgery. A total of 48 different treatment procedures have been separately tested to discover the public sentiment for each of the treatment.

TABLE 3. PERFORMANCEMEASURES

Measure	Value (%)
Accuracy	85.73
Misclassification Rate	14.2
True Positive Rate (recall /sensitivity)	92.1
False Positive Rate	60.8
Specificity	39.1
Precision	91.7
Prevalence	88.1
F-Score	91.9

TABLE 2. NAÏVEBAYESCONTINGENCY

N=1886		Actual		
		Positive	Negative	
Predicted	Positive	1529(True Positive)	137(False Positive)	1666(Predicted positive)
	Negative	132(False Negative)	88(True Negative)	220(Predicted negative)
		1661(Actual positive)	225(Actual negative)	

The obtained results are quite interesting, which are thoroughly discussed in the following section.

4. RESULTS

The results are visualized by simple histograms representing frequency of tweets with respect to their polarities. Polarity strength ranges from -5 to +5 (i.e., very negative to very positive), where as zero is considered as neutral, being neglected in sentiment analysis because the goal of this study is to get polarities. Experiments have been conducted on each plastic surgery procedures. However, for brevity some results are reported as representation. In particular, results about treatments including body countering, chin surgery, and laser skin resurfacing are reported. The comparative results for arm lift, body lift, neck lift, Espanol, tummy tucks, fat transfer and permanent makeup are also presented.

Fig. 2 shows that the body contouring have both positive as well as negative sentiments. Tweets having positive scores (i.e. 1 and 2) are greater in number, while negatively classified tweets are 66.67% for score -1, 8.33% for each -3 and -4 score. Therefore, largely is observed that body contouring has bit more positive public feelings.

Similarly in Fig. 3, chin surgery has been observed with both positive and negative polarities ranging between -3 to 2. These results represent negative disposed regarding the chin surgery procedure.

Similarly, in Fig. 4 negative sentiments are observed. These results reveal the fact that generally cheek augmentation is not publically considered as a good surgery.

Liposuction procedure has more negative public opinion as Fig. 5 clearly depicts it.

Fig. 6 compares the public perspective for 6 different plastic surgery treatments in general. Precisely, body lift and body contouring both have similar public attitude

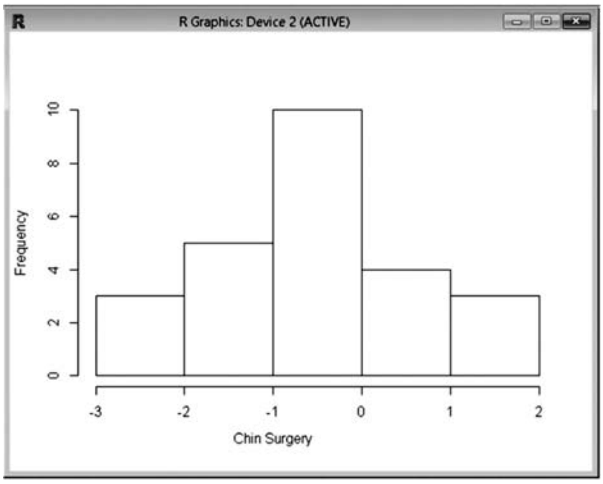

FIG. 3. HISTOGRAM FOR CHIN SURGERY

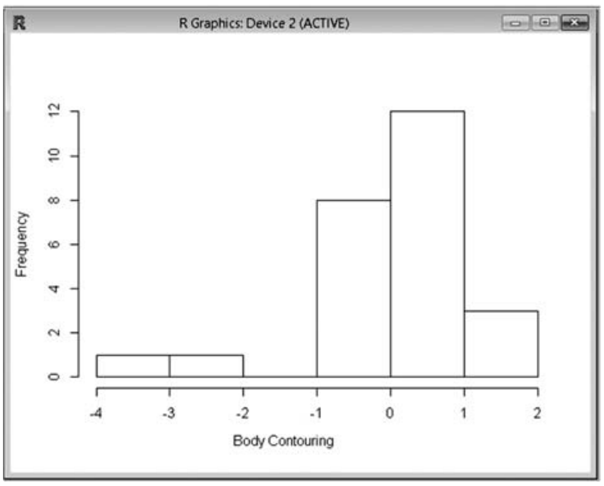

FIG. 2. HISTOGRAM FOR BODY CONTOURING

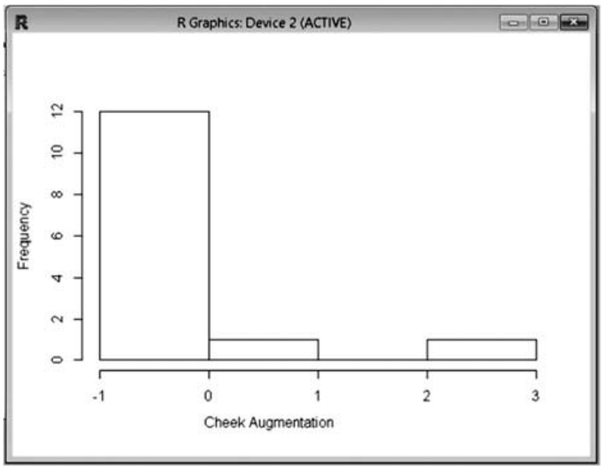

FIG. 4. HISTOGRAM FOR CHEEK AUGMENTATION

that is quite negative. The mostly negative sentiments are for Arm Lift as compared to other 5 treatment procedures. The treatment Espanol has almost neutral public partiality. Likewise, Neck Lift is observed having negative opinions. However, Tummy tuck has been found having both positive and negative polarities but negativity dominates.

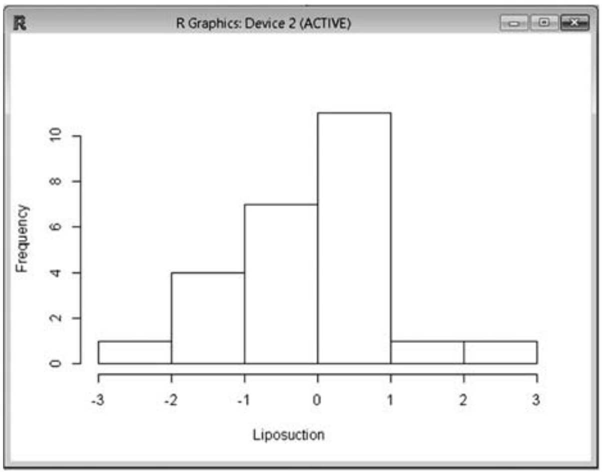

FIG. 5. HISTOGRAM FOR LIPOSUCTION

Skin Cancer Removal procedure has only neutral polarity as shown in Fig. 7. Since, Skin Cancer is a serious disease, public might not have good or bad views about its treatment. Additionally, Fat Transfer and Ear Surgery both have more negative inclination as compared to Permanent Makeup. Contrary, Dermabrasion is observed with positive inclinations.

Largely summarized results are illustrated in Fig. 8, which represents plastic surgery treatment having positive attitude. These results are coherent with the report 2013 of American Society of Plastic Surgeons that a certain variable percentage is increased in plastic surgery trend yearly.

5. CONCLUSION

The analyzed results show the fact that plastic surgery treatment is considered having both polarities: positive as well as negative with respect to the public opinions. However, summarized results represent the willingness of

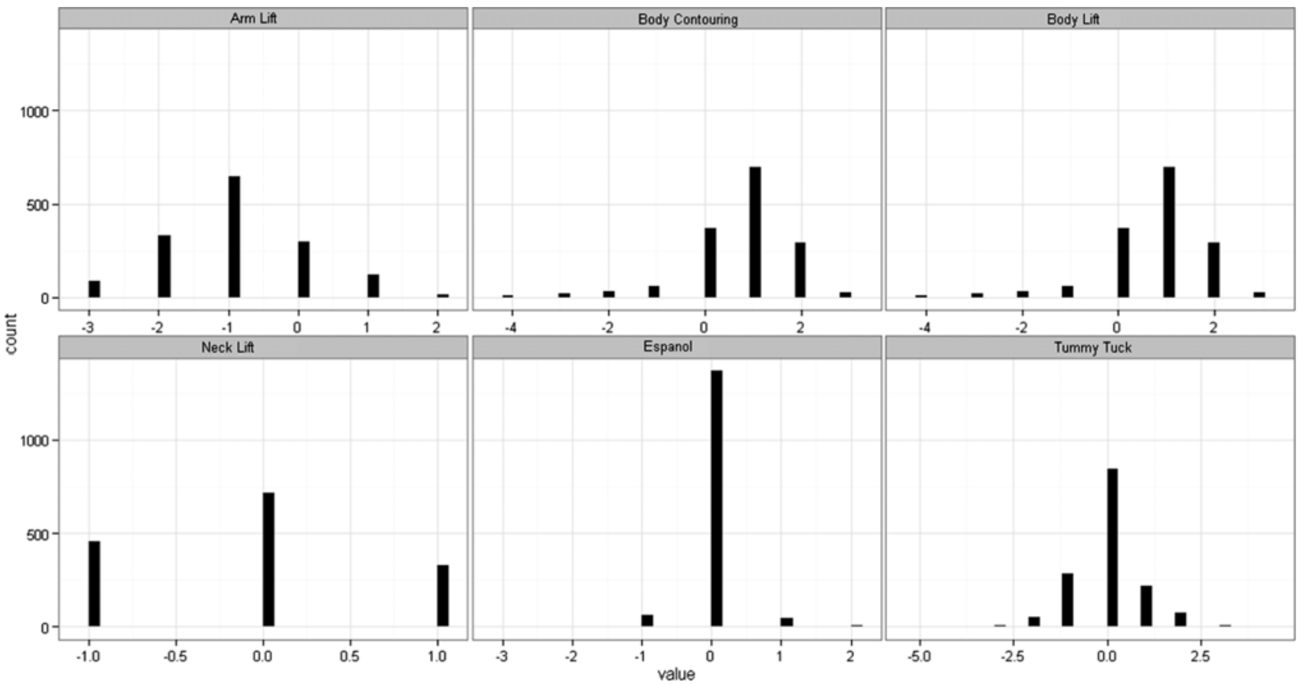

FIG. 6. COMPARATIVE HISTOGRAMS-1

people to undergo plastic surgery treatments to change their appearance despite of its side effects. The proposed approach presented an effective way to derive the individual's opinions from tweets about surgery treatments, which affects human look and careful decisions. The experimental results show the effectiveness of the proposed approach, which may help potential group of people who wish to have surgery treatments in making better decisions. Moreover, the results also highlight the general perception of Twitter users about plastic surgery treatment. This perception may allow surgery providers to streamline people's concerns.

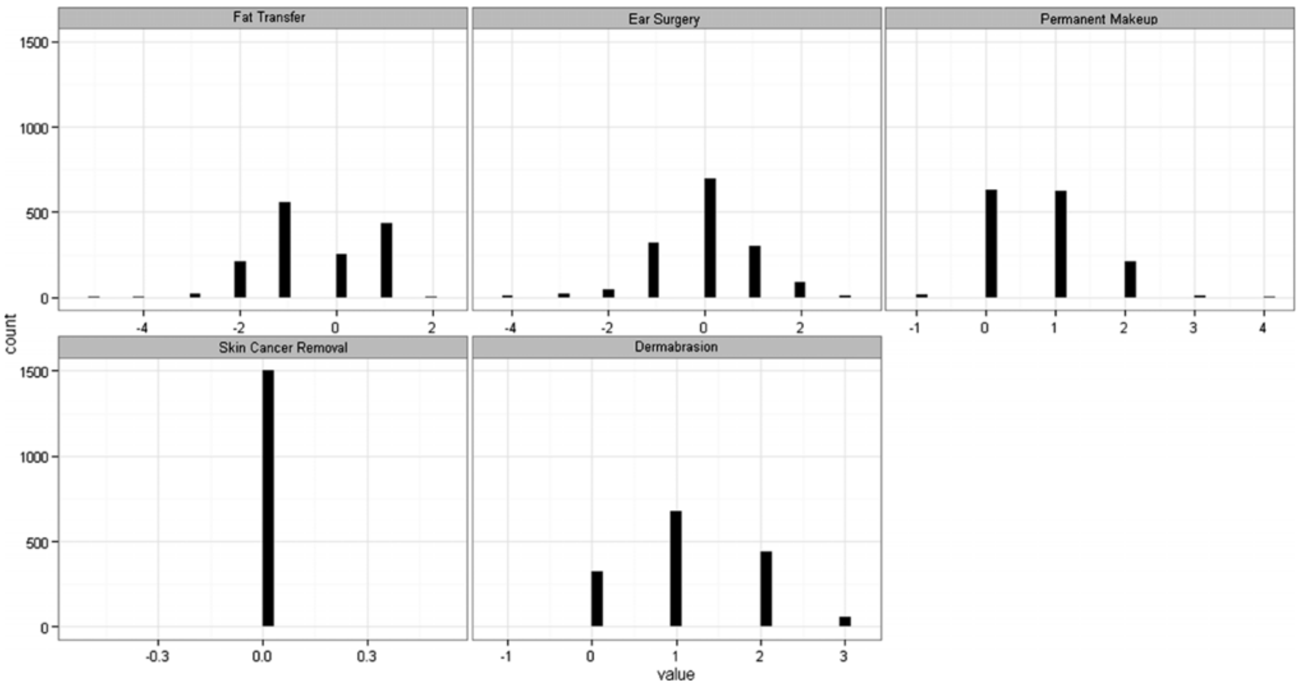

FIG. 7. COMPARATIVE HISTOGRAMS-2

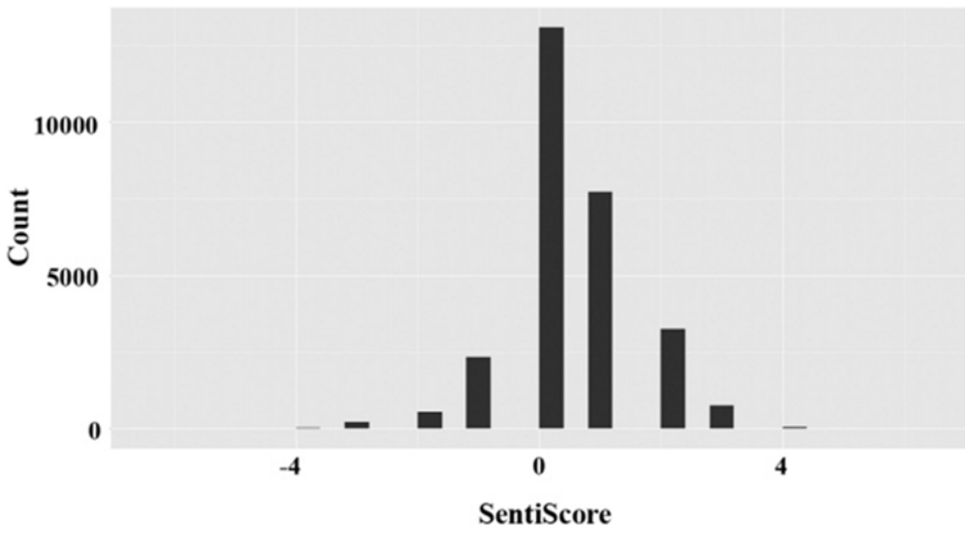

FIG. 8. SUMMARIZED HISTOGRAM

6. FUTURE WORK

As the future work, this research will be extended to grouping the sentiment analysis based ongeo-location, time stamp, gender and age to deeply and concretely discover public opinions regarding different plastic surgery procedures. Furthermore, to extract information from web blogs, social networks and review websiteswith the help of lexicon dictionarieswill be focused in future, which may potentially yield commendable results.

ACKNOLWEDGEMENTS

The authors acknowledge to Prof. Tehseen Hafiz, Department of Software Engineering, for the guidance in the research. The authors are also thankful to Prof. Dr. Bhawani Shankar Chowdry, Dean, Faculty of Electrical, Electroni & Computer Engineering, and Prof. Dr. Mukhtiar Ali Unar, Director, Institute of Information & Communication Technologies, Mehran University of Egnineering & Technology, Jamshor, Pakistan, for their guidance and making IMTIC'15 as a platform for publishing the work and providing research facilities.

REFERENCES

[1] American Society of Plastic Surgeons, "2013 Plastic Surgery Statistics Report", ASPS National Clearing House of Plastic Surgery Procedural Statistics, (Available at http://www.plasticsurgery.org, (Last Accessed 22nd October, 2014).

[2] Pang, B., and Lee, L., "Opinion Mining and Sentiment Analysis", Journal of Foundations and Trends in Information Retrieval, Volume 2, Nos. 1-2, pp. 1-135, USA, 2008.

[3] Tumasjan, A., Sprenger, T.O., Sandner, P.G., and Welpe, I.M., "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment", Proceedings of International Conference on Weblogs and Social Media, Washington, DC, USA, May 23-26, 2010.

[4] Wang, H., Can, D., Kazemzadeh, A., Bar, F., and Narayanan, S., "A System for Real-Time Twitter Sentiment Analysis of 2012 US Presidential Election Cycle", Proceedings of Association of Computational Linguistics System Demonstrations, Jeju Island, Korea, July 10, 2012.

[5] Bollen, J., Mao, H., and Zeng, X.J., "Twitter Mood Predicts the Stock Market", Journal of Computer Science, Volume 2, No. 1, pp. 1-8, 2011.

[6] Mao, H., Counts, S., and Bollen, J., "Predicting Financial Markets: Comparing Survey, News, Twitter and Search Engine Data", Computing Research Repository, Volume ABS/1112.1051, 2011.

[7] Lek, H.H., and Poo, D.C., "Aspect-Based Twitter Sentiment Classification", Proceedings of IEEE International Conference on Tools with Artificial Intelligence, Herndon, VA, USA, November 4-6, 2013.

[8] Ji, X., Chun, S.A., and Geller, J. "Monitoring Public Health Concerns Using Twitter Sentiment Classifications", Proceedings of IEEE International Conference on Healthcare Informatics, Pennsylvania USA, September 9-11, 2013.

[9] Athar, A., "Sentiment Analysis of Scientific Citations", Technical Report Number 856, University of Cambridge, Computer Laboratory, Available online at http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-856.pdf, ISSN:1476-2986.

[10] Beykikhoshk, A., Arandjelovic, O., Phung, D., Venkatesh, S., and Caelli, T., "Data-Mining Twitter and the Autism Spectrum Disorder: A Pilot Study", Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, China, August 17-20, 2014.

[11] Ofek, N., Caragea, C, Rokach, L., Biyani, P., Mitra, P., Yen, J., Portier, K., Greer, G., "Improving Sentiment Analysis in an Online Cancer Survivor Community Using Dynamic Sentiment Lexicon", Proceedings of International Conference on Social Intelligence and Technology, USA, May 8-10, 2013.

[12] Hodeghatta, U.R., "Sentiment Analysis of Hollywood Movies on Twitter", Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Canada, August 25-28, 2013.

[13] Mahmood, T., Iqbal, T., Amin, F., Lohanna, W., and Mustafa, A., "Mining Twitter Big Data to Predict 2013 Pakistan Election Winner", Proceedings of 16th International Multi Topic Conference, Pakistan, December 19-20, 2013.

[14] Karamibekr, M., and Ghorbani, A., "Sentiment Analysis of Social Issues", Proceedings of IEEE International Conference on Social Informatics, Switzerland, December 14-16 2012.

[15] Claster, W.B., Cooper, M., and Sallis, P., "Thailand-Tourism and Conflict: Modeling Sentiment from Twitter Tweets Using Naïve Bayes and Unsupervised Artificial Neural Nets", Proceedings of IEEE 2nd International Conference on Computational Intelligence, Modelling and Simulation, Indonesia,September 28-30, 2010.

[16] Twitter Developer's Documentation, available online at https://dev.twitter.com/docs, [last accessed 2/10/2014].

[17] Hu, M., and Liu, B., "Mining and Summarizing Customer Reviews", Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, Washington, USA, August, 22-25, 2004.

[18] Breen J., "R Tutorial: Mining Twitter for Consumer Sentiment", Online Tutorial Available and Resources available at, http://www.jeffreybreen.com/code, [last accessed 01-08- 2014].

[19] Manning, C.D., Raghavan, P., and Schütze, H., "Introduction to Information Retrieval", Cambridge University Press, Volume 1, [ISBN: 0521865719], 2008.